DOCUMENT RESUME

ED 234 611 FL 013 929

AUTHOR O'Brien, Maryann
TITLE The Classroom Teacher and the Case Study: The Perils and Possibilities of Conducting Your Own Research.
PUB DATE 83
NOTE 8p.; Paper presented at the Annual Convention of Teachers of English to Speakers of Other Languages (17th, Toronto, Ontario, Canada, March 15-20, 1983).
PUB TYPE Viewpoints (120) -- Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Case Studies; Educational Research; *English (Second Language); Higher Education; Language Research; Language Teachers; *Research Methodology; *Research Problems; Second Language Instruction; Second Language Learning; Writing Instruction

ABSTRACT

The problems of classroom teacher-conducted research are examined with reference to a particular case study. The case study involved a field test of "A Practical Guide for Advanced Writers in ESL" over a semester in two sections of an intermediate English as a second language (ESL) writing class for college-bound students at the University of Houston. Both the experimental and control sections kept error charts emphasizing spelling, diction, and syntax. T-unit analysis was used to assess the relative gains of the experimental versus the control sections. A major difficulty with the study was attrition, since 50 percent of one of the sections was lost during the course of the research. Results showed no difference between the control and experimental groups, although both registered substantial improvements. In order to further investigate the nature of these improvements it would be necessary to have a control group of students who did not receive writing instruction. However, it is impossible to justify teaching an intermediate ESL class of university-bound students without writing instruction. This is one of many limitations on teacher-conducted research. A final difficulty is the tendency of a teacher to hope for a certain set of results. (RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Maryann O'Brien

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED234611

FL013929

The classroom teacher and the case study: the perils and possibilities of conducting your own research.

Maryann O'Brien, Curriculum Coordinator, Language and Culture Center, University of Houston.

Empirical research studies in first and second language acquisition can give teachers valuable insights into the learning process, but many of us want to carry out our own "research" in the classroom to find the best methods of day by day instruction. Although these teacher-generated case studies may be well designed, there are many potential sources of error we must be aware of before we draw conclusions. Some of these are questions of N-size, selection of subjects, mortality, measures of evaluation, and statistical procedures. These problems are illustrated by a case study conducted over a semester in two sections of an intermediate ESL writing class at the Language and Culture Center, University of Houston. The study will first be described, and then the problems involved in conducting one's own research will be addressed.

During the past summer semester, I field tested A Practical Guide for Advanced Writers in ESL by Paul Munsell and Martha Clough (Macmillan - in press.) The assumption which the book is based on is that the best way to learn to write is by writing. The authors feel that a sense of fluency should come before too much rhetorical discussion of organization, coherence, and the like. The students are expected to write in class every day for about twenty minutes. The instructor reads and marks the compositions with the standard correction symbols. Students correct their mistakes the following day, but they do not rewrite the entire composition. I decided to design a case study to try to find out how thoroughly papers had to be marked. I wanted to see whether I could save time and help students learn to monitor their writing at the same time. I was also curious about several other questions: how much could students assimilate from these corrections; would they look up the grammar explanations for the mistakes they made; would they remember the correct form on their next paper. The

goal of every writing teacher is to make students independent of teacher controls. Students are going to the university where they have to compete with native speakers and writers and their professors either will not or cannot take the time to separate their English writing ability from the content of their papers.[1]

I was teaching two sections of an upper intermediate writing class. For the case study, I used one class as a control group and the other as an experimental group. I flipped a coin to determine the groups because I wanted the choice random. For the control group, I marked every error in syntax, spelling, and diction with the correction symbols. Students kept error charts and corrected their mistakes every day. For the experimental group, however, I wanted to see if they could assimilate and improve more quickly when they had only one or two points to concentrate on, so I limited the marks on their papers in the following ways: I felt that spelling and diction errors should always be noted. Spelling and vocabulary are two areas that need constant reinforcement, good writing being more than simply a matter of correct grammar. As for syntax and grammar errors, I decided to focus on one each week. The first week, for example, I marked punctuation and capitalization errors, the second week tense mistakes, and so on. Errors marked were not cumulative. The second week, for example, I ignored punctuation and capitalization misuse. Before the week's writing assignments began, these points were not specifically taught because it would not have been fair to the control group, but students in the experimental group were told which "point" to concentrate on for that week. Both groups could, of course, ask for help as they were writing. The experimental group also kept error charts, but they were limited to three divisions: spelling, diction, and the focal point of syntax for that week.

The next problem was the evaluation of the results at the end of the semester. I could not simply count the errors marked on the papers because the control group naturally had many more. And there would be no point in my counting mistakes for the experimental group that I hadn't been marking all term since one of the things I was looking for was a way to reduce grading time for the instructor. If marking only three types of errors

helped students improve, or even if there were no difference between their improvement and that of the other group, then it should not be necessary to take the instructor's time to note each single mistake. I needed an objective means of determining student progress. I decided to use T-unit analysis. To get an indication of student progress after the course ended, I took the first composition written by both groups at the beginning of the term before the experiment started, and counted the total number of T-units in each paper. Then I counted error-free T-units and found the ratio of error-free to total T-units. At the end of the semester, I did the same type of count with the last composition written by both groups, so that I had, in effect, a pretest and a posttest. What I hoped for was that the experimental group would have improved more than the control group.

Before discussing the results, a few words about T-unit analysis are in order. Naturally there are criticisms of this method as an index to good second language writing.It was developed as a way to judge syntactic complexity in first language acquisition. One problem with it is that it naturally works better for more advanced learners. My upper intermediate class may not have been advanced enough for this kind of analysis, even though this is the level at which many of them leave intensive English programs and try to enter American universities. The ratios of error-free to total T-units I found were very low, which might be an indication that the general language competency level of these groups was not high enough. Another criticism of T-unit analysis is that it does not consider stylistic appropriateness, but by using "diction" as a possible error type, I tried to take this somewhat into account. Nor does it help the student who is trying out new constructions before they are perfectly mastered. But I needed some objective measure of evaluation, and T-unit analysis, while not perfect for ESL, is at least accepted as long as exaggerated claims are not made based on its use alone.

I want to mention here several problems with this study. These refer to the "perils" of my title. I mentioned above that N-size, selection of subjects, mortality, measures of evaluation and statistical procedures were potential problems. Together with the particular concern with the measure

of evaluation used, mortality was the biggest problem. By the end of the semester, I had lost exactly half of one class, reducing the N-size, which began the same for each class, by 50% in one section. Correlations were done between the two Michigan tests and the sample compositions that both groups wrote - one as a pretest and one as a posttest, but naturally, those students who left early had to be dropped from the study and their pretests could not be used when figuring results. Selection of subjects may also have been a problem. I mentioned above that I had flipped a coin to determine which section would be the experimental group, but there was no randomization of individuals to each group. I had to take the two sections as they came to me, and I had to deal with them as two complete groups throughout the semester. In Foundations of Behavioral Research, Fred Kerlinger writes that such a design is subject to the weaknesses due to the possible lack of equivalence between the groups: "The experimental treatments should be assigned at random. Then the similarity of the groups should be checked using any information available - sex, age, etc. . . . Although one cannot have the assurance that randomization gives, if these items all check, one can go ahead with the study knowing at least that there is no evidence against the equivalence assumption." (Kerlinger 1964) If we could take students out of their assigned classes and regroup them for the experiment, this would be possible, but in classroom studies that cover an entire semester of treatment, classes have to be left intact. But as we ran a T-test on the pretests, and as there was no significant difference between the two means on the Michigan, we can assume that as far as what the Michigan measures, the students were from the same population. Both groups were equally divided as to sex, but language and country backgrounds were different. These are variables which I did not take into account.

Another "peril," at least for me, and no doubt for many classroom teachers, is statistical procedures. Fortunately, our program has its own systems analyst, and I was able to get help and advice in collecting data and determining results. Making use of the computer is definitely one of the

"possibilities" of my title. Even if your program does not have its own computer specialist, most universities have statistics or computer labs where help is available. If so, I would encourage their use. It saves a lot of time, and allows us to feel more confident about the results.

In this case study, we found only one significant result, and that was in the area of gains. There was no correlation between the Michigan exams and the sample compositions, but we can't really tell what that indicates. It could mean that the two instruments were measuring different things, or it could mean that the T-unit analysis I used was not appropriate for this level, as I mentioned before, or it could mean something else. Even the Michigan gain was not significant. But in comparing the sample essays, pre and post, we found a twenty-two point gain for both groups, which was significant at the .001 level of our T-test. This does not show that one group turned out to be better writers than the other group, but it shows that both groups improved significantly in their writing ability. I interpreted these results to mean that whatever I did to the students didn't seem to harm them. If I marked only three error types, they gained. If I marked all their errors throughout the semester, they gained the same amount. Perhaps just by writing, they became better writers, which would support the author's premise that students learn to write by writing. But there are other "perils" to consider here. The only way to be certain of this would be to have had a third group all semester who studied no writing at all. If they had gained as much as the other two groups, then it might not be the practising of writing so much as general exposure to English over a semester. But then I would also have needed a fourth group who were not intensive English students at all, but who were just non-native speakers of English in the community. Even if I had this third group at my disposal, there are practical, humanitarian concerns to be considered. Could we justify teaching an intermediate ESL class of university bound students without doing any writing? Our students are not 'guinea pigs,' but tuition-paying clients, so there have to be limitations on what we do to them. If we want to conduct a little study that takes a sample from

a whole program for a few hours, then we might be less restricted, but one matter we have to consider when planning any design that lasts an entire semester is whether we are neglecting anything essential to one group or the other.

The first thing to do when you finish a case study is to decide what you would do differently next time. I found many "perils" that I would wish to avoid. Specifically, I would try to use more subjects, but where do we get them? While an N-size of six seems rather ridiculous, it is important to mention it in any study. In the past, it was not unknown for extremely small N-sizes to appear in fine print at the bottom of published ESL research. I would also do the study in the fall or spring terms when the attrition rate is not so high. It would also be interesting to try other approaches to teaching writing to see whether that would make a difference. I might, for example, include more grammar work. (In our program, the grammar component is often taught by a different instructor.) Finally, I would look for other objective measures of evaluation, or I would use T-unit analysis with a more advanced class. But even if T-unit analysis was not appropriate for this level, I could tell by looking at students' papers that there was quite a big improvement from the beginning of the term to the end, and I could also tell that there was not that much difference between the two groups' writing ability. Our subjective judgement as ESL instructors has been exercised for years, and I do think it still has its place.

One final "peril" involved in conducting our own research is that we must be careful not to hope for particular results before we begin, as I did. We don't want to prejudice the outcome during the course of treatment. We won't always find what we're looking for, we often find something we weren't looking for, and very often we find nothing at all. Studies such as these don't answer questions so much as raise more questions. The purpose of any research, whether informally carried out by the classroom teacher, or carefully controlled by experts, is investigation. This does

not mean that the classroom teacher should not attempt some of these experiments. As long as we are aware that what we're doing in cases like these is trying to get some indication of what seems to work best, I believe we should continue. The important thing is to quantify our impressions and make decisions based on verifiable results instead of basing all our teaching decisions on intuition. We learn from the "perils," but the "possibilities" are far more important.

[1]We conducted surveys of the professors in foreign students' major fields here last year, and found that the majority of the professors found grammar the most serious problem in foreign student writing, which was not surprising. As ESL teachers, we can usually find the meaning behind their garbled sentences by second nature, but we can't expect others to take the time to do this.

References:

Gaies, Stephen J. 1980. T-unit analysis in second language research: applications, problems and limitations. TESOL Quarterly 14:53-60.

Kerlinger, Fred N. 1964. Foundations of behavioral research. New York: Holt, Rinehart and Winston.

Munsell, Paul and Martha Clough. A practical guide for advanced writers in English as a second language. (in press) New York: Macmillan.

Perkins, Kyle. 1980. Using objective methods of attained writing proficiency to discriminate among holistic evaluations. TESOL Quarterly. 14: 61-69.